

Taxonomic Status of the *Bemisia tabaci* Complex (Hemiptera: Aleyrodidae) and Reassessment of the Number of Its Constituent Species

Wonhoon Lee¹, Jongsun Park², Gwan-Seok Lee³, Seunghwan Lee^{4*}, Shin-ichi Akimoto^{1*}

¹ Laboratory of Systematic Entomology, Department of Ecology and Systematics, Graduate School of Agriculture, Hokkaido University, Kita-ku, Sapporo, Japan,

² Stronghold Innovation Center, Gocheok-dong, Guro-gu, Seoul, Korea, ³ Crop Protection Division, National Academy of Agricultural Science, RDA, Gyonggi-do, Korea,

⁴ Insect Biosystematics Laboratory, Research Institute for Agricultural and Life Sciences, Seoul National University, Seoul, Korea

Abstract

Bemisia tabaci (Hemiptera: Aleyrodidae) is one of the most important insect pests in the world. In the present study, the taxonomic status of *B. tabaci* and the number of species composing the *B. tabaci* complex were determined based on 1059 *COI* sequences of *B. tabaci* and 509 *COI* sequences of 153 hemipteran species. The genetic divergence within *B. tabaci* was conspicuously higher (on average, 11.1%) than interspecific genetic divergence within the respective genera of the 153 species (on average, 6.5%). This result indicates that *B. tabaci* is composed of multiple species that may belong to different genera or subfamilies. A phylogenetic tree constructed based on 212 *COI* sequences without duplications revealed that the *B. tabaci* complex is composed of a total of 31 putative species, including a new species, *JpL*. However, genetic divergence within six species (*Asia II 1*, *Asia II 7*, *Australia*, *Mediterranean*, *New World*, and *Sub Saharan Africa 1*) was higher than 3.5%, which has been used as a threshold of species boundaries within the *B. tabaci* complex. These results suggest that it is necessary to increase the threshold for species boundaries up to 4% to distinguish the constituent species in the *B. tabaci* complex.

Citation: Lee W, Park J, Lee G-S, Lee S, Akimoto S-i (2013) Taxonomic Status of the *Bemisia tabaci* Complex (Hemiptera: Aleyrodidae) and Reassessment of the Number of Its Constituent Species. PLoS ONE 8(5): e63817. doi:10.1371/journal.pone.0063817

Editor: William J. Etges, University of Arkansas, United States of America

Received: October 29, 2012; **Accepted:** April 6, 2013; **Published:** May 13, 2013

Copyright: © 2013 Lee et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was supported by a Grant-in-Aid for Scientific Research (No. P.11079 for Postdoctoral Fellowship for Foreign Researchers and No. 23370037 to SA) from the Japan Society for the Promotion of Science. This research was also supported by a grant from the Agricultural R&D Promotion Center, Korea. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: seung@snu.ac.kr (SL); akimoto@res.agr.hokudai.ac.jp (SA)

Introduction

The sweet potato whitefly, *Bemisia tabaci* (Gennadius, 1889) (Hemiptera: Aleyrodidae), is one of the most important agricultural pests in the world [1–3]. This species attacks several host plants including ornamental and vegetable plants, grain legumes, and cotton by ingesting the phloem sap [4] and transmits begomoviruses (Geminiviridae) [5,6], inducing a large amount of economic damage [3]. *B. tabaci* is a complex of at least 24 morphologically indistinguishable species [3]. These species have been mainly distinguished by molecular methods because of difficulties in detecting morphological and biological differences among the species [3,7,8]. Since the 1990s, various molecular methods such as allozymes, random amplified polymorphic DNA, microsatellites, mitochondrial genes, and nuclear genes have been used for defining several groups in this complex [3]. Currently, mitochondrial *cytochrome oxidase subunit I* (*COI*) gene is primarily used to define several species of the *B. tabaci* complex [9].

Until 2010, large genetic divergence within *B. tabaci* had posed the question of whether *B. tabaci* was one biological species composed of several biotypes or a complex of biological species [10]. Dinsdale et al. [9] addressed this problem and reported that 24 constituent species of *B. tabaci* were distinguished by a threshold of 3.5% difference based on 454 *COI* sequences and biological characters. As a result, many researchers have distinguished the

constituent species in the *B. tabaci* complex based on the 3.5% genetic boundary [11–13]. However, the 3.5% boundary has not been compared with species boundaries of other groups regarding the same position of *COI* sequence. Thus, it is required to confirm the ground for the 3.5% boundary in the *B. tabaci* complex using much more sequence data. Until now, it is still unresolved whether the *B. tabaci* complex is monophyletic, consisting of cryptic species, or is polyphyletic, including species of different genera, subfamilies, and families. This question is raised because genetic divergence in *COI* sequences within the *B. tabaci* complex (range, 0–34%) [9] is much higher than that observed between congeneric species in other hemipteran groups such as aphids (range, 0–11%) [14].

Recently, new species of the *B. tabaci* complex have been reported based on the 3.5% boundary. For example, Hu et al. [13] reported four new species (*Asia II 9*, *Asia II 10*, *Asia III*, and *China 3*), and then two new species (*Asia I-India* and *New World 2*) were later reported by Chowda-Reddy et al. [11] and Alemandri et al. [12], respectively. As a result, a total of 30 species have been reported in the *B. tabaci* complex. However, this estimate is still uncertain because not all of the *COI* sequences of *B. tabaci* were included in these studies (i.e., Dinsdale et al. [9], Chowda-Reddy et al. [11], Alemandri et al. [12], and Hu et al. [13]). For example, Ueda et al. [15] reported one putative new species, *JpL*, of the *B. tabaci* complex; however, *COI* sequences of the *JpL* species have

not been considered in the subsequent calculations of the number of species in the *B. tabaci* complex [9].

A number of species concepts have been proposed to date, and the usefulness and applicability of respective species concepts has been critically discussed [16]. A species concept based on genetic distance has often been criticized because genetic distance does not always reflect the existence of pre- or post-mating isolation [17]. This criterion of the biological species concept (BSC) [18] has been applied to diverse taxonomic groups with different genetic systems and life histories. However, unusually, the delimitations of the species comprising the *B. tabaci* complex have been based on a discontinuous distribution of genetic distances in genetic markers without considering biological data. Currently, most researchers have followed ‘phylogenetic’ species concept [9], in which species boundaries are constructed with phylogenetic analyses and pairwise comparisons of genetic distances. Thus, in the present study, we consistently use “species” as a group that is discontinuous from other such groups in terms of genetic distance.

In this study, first, we examined the taxonomic status of *B. tabaci* by calculating genetic divergences among a larger number of *COI* sequences collected from databases. As a comparison, genetic divergence among 509 *COI* sequences of 153 hemipteran species was calculated according to three taxonomic levels (within species, between congeneric species, and between genera in the same family), and the results were compared with genetic divergence among 1059 *COI* sequences of *B. tabaci*. Second, a phylogenetic tree was constructed based on 212 *COI* sequences of *B. tabaci* and reassessed the number of species composing the *B. tabaci* complex. In this re-analysis, we revised the threshold genetic distance that most appropriately represents species boundaries within the *B. tabaci* complex by observing the distribution of genetic divergences among the *COI* sequences of the *B. tabaci* complex. As a control group to the *B. tabaci* complex, genetic divergence of 34 aphid species was calculated according the three taxonomic levels. Third, to examine the advantage and characteristics of the *COI* gene as a genetic marker, we compared patterns of genetic divergences among the *COI* gene and ten other mitochondrial genes by using 20 hemipteran mitochondrial genomes.

Results

Genetic Divergence at Three Taxonomic Levels

In the *B. tabaci* dataset, genetic divergence was detected among the *COI* sequences, ranging from 0% to 24.1% with an average of 11.1% (Table 1). Genetic divergence among *B. tabaci* individuals was categorized into three separate groups based on their frequency distributions (Figure 1A): group 1 (divergence ranging from 0–4%), group 2 (range, 4–11%), and group 3 (range, 11–25%). Dinsdale et al. [9] reported that the genetic divergences among 198 *COI* sequences of *B. tabaci* and four sequences of *B. triplex*, *B. afer*, and *B. subdecipiens* ranged from 0% to 34%; the range was subsequently separated into four groups: group 1 (range, 0–3%), group 2 (range, 3.5–11%), group 3 (range, 11–26%), and group 4 (range, 26–34%). Because our analysis excluded the four sequences of *B. triplex*, *B. afer*, and *B. subdecipiens*, the group 4 in Dinsdale et al. [9] was not used.

The hemipteran species dataset exhibited a different tendency from the *B. tabaci* dataset in the average genetic divergence within species (Table 1); the average genetic divergence increased as the taxonomic level rose (Figure 1B). The combined dataset (*B. tabaci*+hemipteran species) also exhibited variation in genetic divergence at the three taxonomic levels (Table 1). However, in

this case, they did not increase with rising taxonomic levels (Figure 1C), and the average genetic divergence among conspecific individuals (10.7%; range, 0–24.1%) was higher than the average genetic divergence between congeneric species (6.5%; range, 0–28.1%).

As a control to the *B. tabaci* complex, we used aphids which are closely related to the genus *Bemisia* in Hemiptera and quantified the pattern of genetic divergences using 47 *COI* sequences of 34 aphid species. The aphid group exhibited average genetic divergences of 0.3% (range, 0–1.5%) among individuals of one species, 5.7% (range, 0.1–11.7%) between different species belonging to one genus, and 12.8% (range, 5.2–18.9%) between different genera belonging to one subfamily.

Defining the Number of Species of the *B. tabaci* Complex

The phylogenetic tree constructed from the 222 *COI* sequences of *B. tabaci*+ten other aleyrodide species revealed that the *B. tabaci* sequences were separated into 31 groups (Figure 2). These groups were named based on five previous studies [9,11–13,15]. The phylogenetic tree was similar in topology to those of the previously reported phylogenetic trees [3,13]. For example, four species (*Indian Ocean*, *Mediterranean*, *Middle East Asia Minor 1*, and *Middle East Asia Minor 2*) were clustered with 99% posterior probability. In addition, eight species (*Asia II 1*, *Asia II 2*, *Asia II 3*, *Asia II 4*, *Asia II 5*, *Asia II 6*, *Asia II 7*, and *Asia II 8*) and seven species (*Asia I*, *Asia III*, *Australia/Indonesia*, *Australia*, *China 1*, *China 2*, and *China 3*) were clustered with 97% posterior probability, respectively. However, some species showed different relationships from the previous studies. Here, *New World* was clustered with four species (*Indian Ocean*, *Mediterranean*, *Middle East Asia Minor 1*, and *Middle East Asia Minor 2*), whereas it was clustered with *Bemisia triplex* in Dinsdale et al. [9]. In addition, *Asia I-India* was clustered with *Asia II 5* with 96% posterior probability, whereas it was clustered with *Asia I* in Chowda-Reddy et al. [11]. Additionally, *New World 2* came into *New World* in the present study; however, it was separated from *New world* in Alemdandri et al. [12]. Among the 31 species, *JpL* has not been reported in other molecular studies conducted since the publication of Ueda et al. [15]. For the first time, we observed that *JpL* species was separated from the other 30 species in the phylogenetic tree.

Figure 3 indicates the first two axes of a canonical discriminant analysis (CDA) for the 31 species of *B. tabaci* based on a matrix of pairwise genetic distances. The first axis in the diagram clearly separates the 31 species assemblages into two groups, with the plots from four species (*Indian Ocean*, *Mediterranean*, *Middle East Asia Minor 1*, and *Middle East Asia Minor 2*) having negative scores, and the plots from the remaining 27 species all having positive scores. The second axis, coupled with the first axis, successfully separated the sampling units into the 27 species. The validity of this classification was tested with a leave-one-out cross-validation method, in which CDA was conducted using all the sequence data excepting one. Then, we tested into which group the excepted sequence was categorized based on the discriminant criteria, and this test was repeated for every sequence by excluding the sequences one by one. The result of this reclassification revealed that the proportion of correct classification was 100%, indicating that the 31 species are clearly distinct (Table 2). Thus, CDA corroborated the 31 groups revealed by the phylogenetic analysis.

Comparing Intra- and Inter-specific Genetic Divergences

Intraspecific genetic divergence of 24 species with two or more different sequences and the interspecific genetic divergence

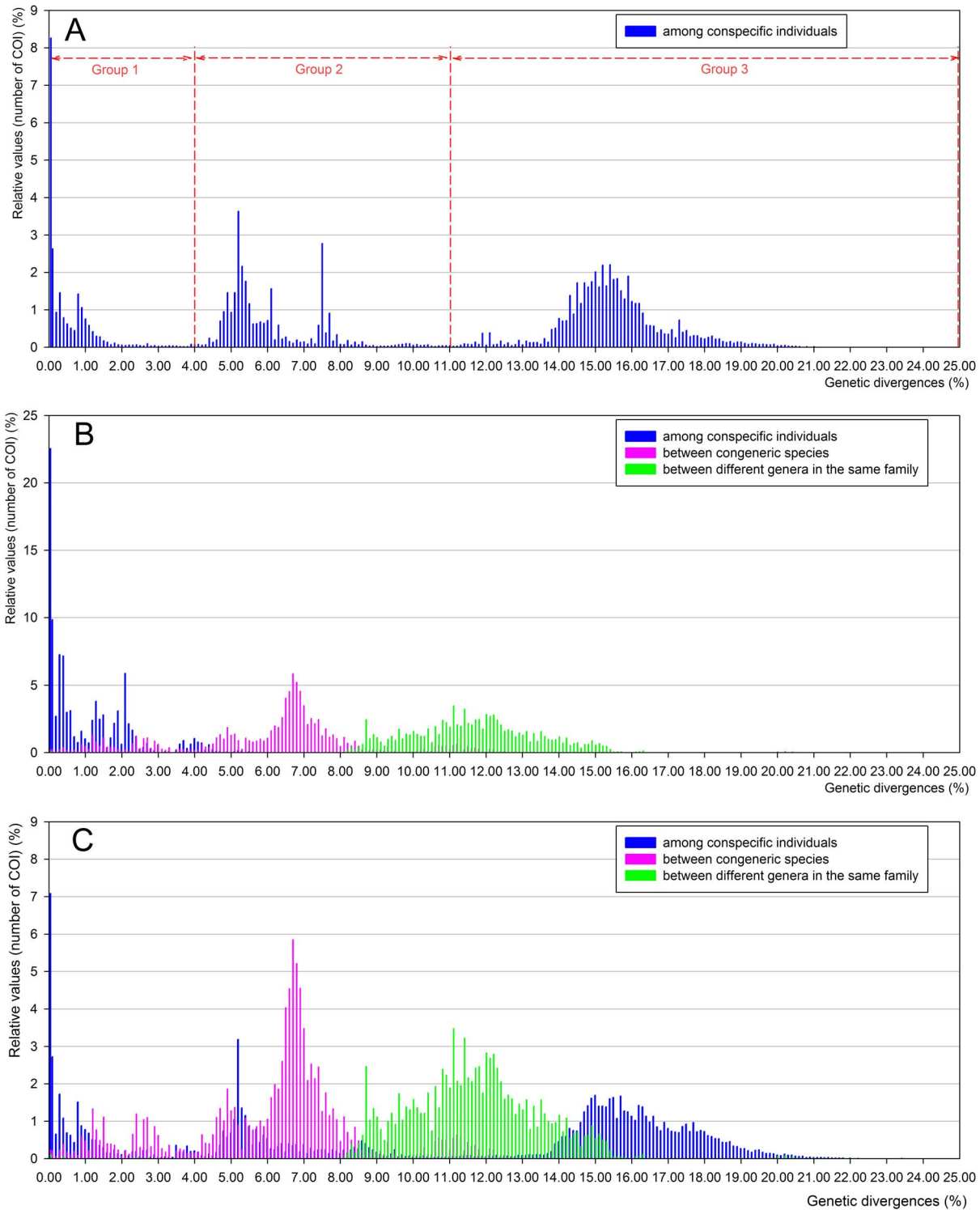


Figure 1. Distribution of genetic divergence estimate for COI sequences according to taxonomic levels. (A) *Bemisia tabaci* dataset consisting of 1059 COI sequences of *B. tabaci*, (B) hemipteran species dataset consisting of 509 COI sequences of 153 hemipteran species, and (C) combined dataset consisting of 1568 COI sequences of 153 hemipteran species and *B. tabaci*. doi:10.1371/journal.pone.0063817.g001

between pairs of closely related species in the *B. tabaci* complex were analyzed using the p -distance and K2P distance models (Figure 4; Tables S1 and S2). Nine pairs of species were selected based on the result of the above-mentioned phylogenetic analysis.

Overall, genetic divergence estimates from the K2P distance model were higher than those from the p -distance model (Tables S1 and S2). In this study, we focused on the results obtained with

Table 1. Genetic divergence estimates of the three datasets at three taxonomic levels.

	Within species			Within genus and between species			Within family and between genera		
	Avg.(%) ^a	Min.(%) ^b –Max.(%) ^c	N.C. ^d	Avg.(%)	Min.(%)–Max.(%)	N.C.	Avg.(%)	Min.(%)–Max.(%)	N.C.
<i>Bemisia tabaci</i>	11.1	0.0–24.1	207046
hemipteran species	1.0	0.0–4.9	6474	6.5	0.0–28.1	12004	12.1	5.8–26.9	27928
<i>B. tabaci</i> +hemipteran species	10.7	0.0–24.1	213171	6.5	0.0–28.1	13629	12.9	5.8–26.9	28566

^aAverage,^bMinimum,^cMaximum, and ^dNumber of comparisons.

doi:10.1371/journal.pone.0063817.t001

the K2P distance model because the 3.5% boundary was calculated based on the K2P distance model [9,19].

Intraspecific genetic divergence varied among the 24 species, ranging from 0.2% to 4.0% (Figure 4; Table S1). Among the 24 species, *JpL* demonstrated the lowest average genetic divergence, 0.4% (range, 0.2–0.7%), whereas *Australia* showed the highest average genetic divergence, 3.8%. Most of the species showed average genetic divergences less than 2.0%, except for two species, *Asia II 6* (with an average of 3.1%) and *Australia* (with an average of 3.8%). However, six species (*Asia II 1*, *Asia II 7*, *Australia*, *Mediterranean*, *New World*, and *Sub Saharan Africa 1*) had the maximum genetic divergences that were higher than the genetic boundary of 3.5%. In particular, *Mediterranean* showed genetic divergences ranging from 0.2% to 4.0%.

Interspecific genetic divergence estimates also varied among the nine pairs of closely related species (Figure 4; Table S2). Among them, *Australia–Australia/Indonesia* showed the largest average genetic divergence, 16.0% (range, 14.1–18.0%), while *China 1–China 2* showed the lowest average divergence, 4.5% (range, 4.2–5.2%). The genetic divergences between all the combinations of the nine pairs were always higher than the 3.5% boundary (Figure 4; Table S2). From 22366 pairwise genetic distances for the 212 *COI* sequences, we observed that the 31 species revealed an average intraspecific genetic divergence of 1.2% (range 0.2–3.9%) and an average interspecific genetic divergence of 15.7% (range, 4.2–24.1%) (Figure 5).

Comparing Genetic Divergences between *COI* and Ten Mitochondrial Genes

A total of 220 sequences of eleven mitochondrial genes were examined by using 20 hemipteran mitochondrial genomes, including *B. tabaci* (NC 006279) (Table S3). Comparisons of the K2P distances among the eleven genes revealed that eight genes (*ATP6*, *COII*, *CytB*, *ND1*, *ND2*, *ND3*, *ND4*, and *ND5*) showed higher pairwise divergences, whereas two genes (*lrRNA* and *srRNA*) showed lower pairwise divergences than did *COI* (Figure 6). Wilcoxon signed rank tests also indicated that *COI* had lower genetic distances among all the species than did the eight genes, but higher genetic distances than did *lrRNA* and *srRNA* (Table 3). For all mitochondrial genes but *COI*, the genetic divergences for each gene exhibited a linear relationship to those for *COI*, suggesting that mitochondrial genes with a higher evolutionary rate could be used as a genetic marker as well as *COI* in the *B. tabaci* complex.

Discussion

Taxonomic Status of the *B. tabaci* Complex

The results of our analyses indicate that the taxonomic status of *B. tabaci* is peculiar, compared to that of other hemipteran species. Among the three datasets examined, the pattern of genetic distances in the hemipteran species dataset (Figure 1B) was similar to the pattern reported in DNA barcode studies with the superclass Hexapoda [20]. Barcode studies have revealed that genetic divergence estimates generally increased from low taxonomic levels (e.g., conspecific individuals and congeneric species) to higher taxonomic levels (e.g., confamilial species) in several orders such as Hemiptera [14], Lepidoptera [21], Diptera [22], and Hymenoptera [23]. However, in the combined dataset (Figure 1C), such an increase in genetic divergences across taxonomic levels was not observed. This result may be mainly due to the broad range of intraspecific genetic divergences observed only in *B. tabaci*. The distribution pattern of genetic divergences in the combined dataset has not been documented in Hemiptera or in other hexapod orders. This result corroborates the possibility that the *B. tabaci* complex consists of several species that may belong to different genera or subfamilies with respect to genetic divergence.

Despite the large genetic divergence, there was no clear evidence for polyphyly in the *B. tabaci* complex. The phylogeny obtained indicated that other *Bemisia* species do not cluster with any species of the *B. tabaci* complex. Although the resolution of *Bemisia* species, including *B. tabaci*, was not sufficient in our phylogenetic tree, it is likely that the constituent species of the *B. tabaci* complex have a separate origin from other *Bemisia* species. Future studies would therefore benefit from a detailed examination as to whether or not the *B. tabaci* complex is paraphyletic to other *Bemisia* species.

The phylogenetic tree and the CDA plots based on the 212 *COI* sequences of *B. tabaci* produced 31 groups, matching the 31 “species” discussed in previous molecular studies. Among these 31 species, 30 have often been reported in several molecular studies conducted since that of Dinsdale et al. [9]. However, the present study is the first to incorporate *JpL* sequences (AB308111, AB308114–AB308117, and AB240697.1) into analysis of the *B. tabaci* complex since Ueda et al. [15]. Samples of the *JpL* species were collected on four host plants in Japan: *Lonicera japonica* (Caprifoliaceae), *Perilla frutescens* (Lamiaceae), *Solanum lycopersicum*, and *Solanum melongena* (Solanaceae) [15]. Because the *JpL* species was separated from the other 30 species of the *B. tabaci* complex in the phylogenetic tree (Figure 2), additional studies are required to determine whether the *JpL* species indeed belongs to the *B. tabaci* complex or constitute a distinct *Bemisia* species.

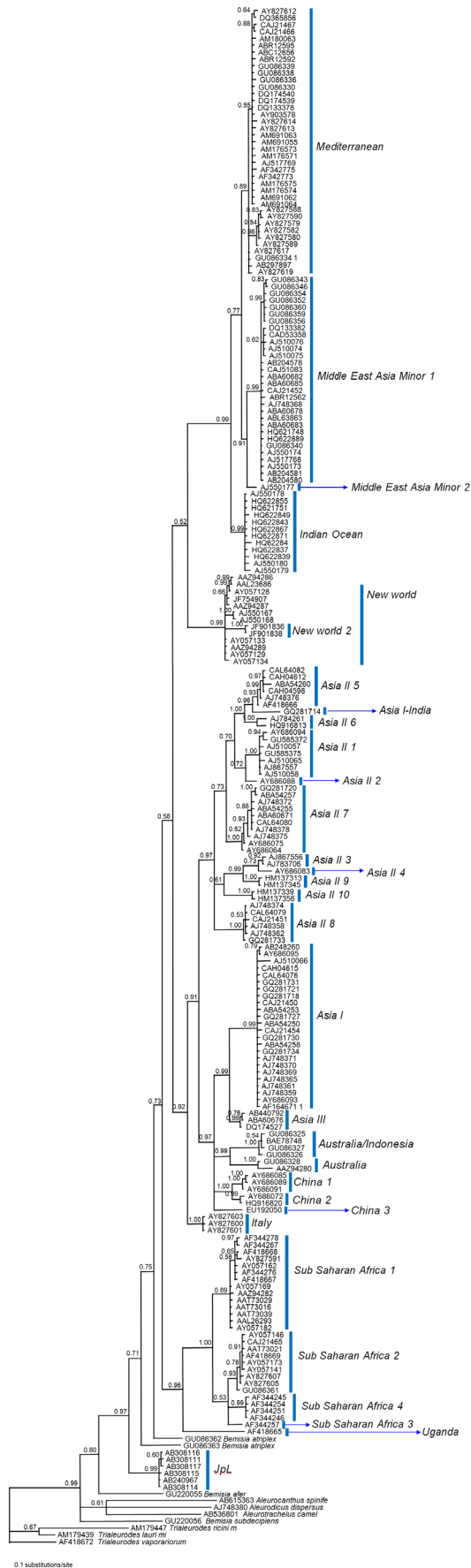


Figure 2. Phylogenetic tree based on 212 *COI* sequences of *Bemisia tabaci* estimated using Bayesian inference analysis. Posterior probability is shown above branches. doi:10.1371/journal.pone.0063817.g002

New Threshold for Species Boundary

Dinsdale et al. [9] suggested the 3.5% boundary for distinguishing 24 species of the *B. tabaci* complex. However, in this study, we observed that estimates of intraspecific genetic divergence were higher than 3.5% for six species (*Asia II 1*, *Asia II 7*, *Australia*, *Mediterranean*, *New world*, and *Sub Saharan Africa 1*) (Figure 4). We also calculated 22366 pairwise genetic distances for the 212 *COI* sequences, and found a gap at 4% in the distribution of genetic divergences (Figure 5). This implies that a genetic divergence of 4% could be a threshold that distinguishes divergences within the same species from divergences between species. Thus, the threshold of the species boundary in the *B. tabaci* complex is required to be replaced with 4.0%.

The difference between 3.5% boundary and 4.0% boundary (in this study) may be due to the addition of *COI* sequences that were reported after or not included in Dinsdale et al. [9]. For example, for *Asia II 7*, intraspecific genetic divergences ranged from 0.3% to 2.9% using five *COI* sequences [9]. However, after adding six more *COI* sequences, the largest genetic divergence increased up to 3.8%. The results of aphid genetic divergences implied that if more sequence data were used and if the pattern of genetic divergences in the *B. tabaci* complex is similar to aphids, the species boundary in the *B. tabaci* complex could increase from 4.0% to 6.0%. This study suggests that the current genetic boundary could vary depending on the sample number of *COI* sequences.

Additional Molecular Makers in the *B. tabaci* Complex

Comparisons of genetic divergences among the 11 mitochondrial genes indicate that respective genes have different patterns of genetic divergences (Table S3; Figure 6). The reason for different levels of genetic divergence among mitochondrial genes has not been clearly explained; however, we confirmed that the *COI* gene exhibits low levels of differentiation compared with the eight genes (*ATP6*, *COII*, *CytB*, *ND1*, *ND2*, *ND3*, *ND4*, and *ND5*) probably due to a low evolutionary rate. Genetic differentiation in the *COI* gene may readily reflect reproductive isolation that occurred a few million years ago [24] without saturation of nucleotide substitutions because of a low evolutionary rate.

Recently, DNA barcode studies have frequently reported that when only *COI* sequences were used, they did not show proper phylogenetic relationships [14,25,26]. This indicates that the use of additional molecular markers is required to resolve phylogenetic relationships in the *B. tabaci*. Incorporation of these additional molecular markers will clarify not only the phylogenetic relationships among the species comprising the *B. tabaci* complex but also the taxonomic status of this complex.

Conclusion

In this study, we confirm that the *B. tabaci* complex is composed of 31 distinct groups or “species” and that the genetic divergence between some species corresponded to that found between different genera or subfamilies. In addition, we observe that the 4.0% genetic boundary was more realistic than the 3.5% in distinguishing species in the *B. tabaci* complex. However, it is not confirmed whether the *JpL* species belongs to the *B. tabaci* complex or not.

Until now, in the *B. tabaci* complex, morphological variation has not been emphasized in spite of large genetic variation among

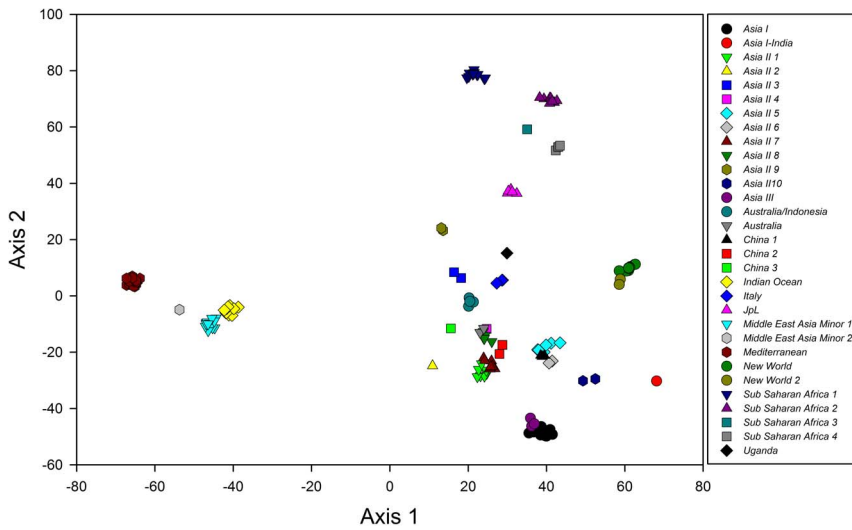


Figure 3. Results of the canonical discriminant analysis based on the 22336 pairwise genetic distances for demonstrating the relationship among the 31 species belonging to the *Bemisia tabaci* complex.
doi:10.1371/journal.pone.0063817.g003

several species [27]. However, in the phylogenetic tree, *B. atriplex* that was distinguished from *B. tabaci* based on morphological characters [28] was clustered with some lineages of *B. tabaci*. Thus, it is necessary to detect morphological differences among several species groups of *B. tabaci*. Although the present study provides useful genetic information in understanding the *B. tabaci* complex, morphology and reproductive incompatibility tests must be essential for applying the biological species concept to the *B. tabaci* complex.

Materials and Methods

Selecting *COI* Sequences of *Bemisia tabaci* and Hemipteran Species

We obtained 15465 *COI* sequences in the Hemiptera from the GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>) by using two keywords, Hemiptera and *COI*. To select 650 bp *COI* sequences located in the 3' region from the 15465 *COI* sequences, alignments

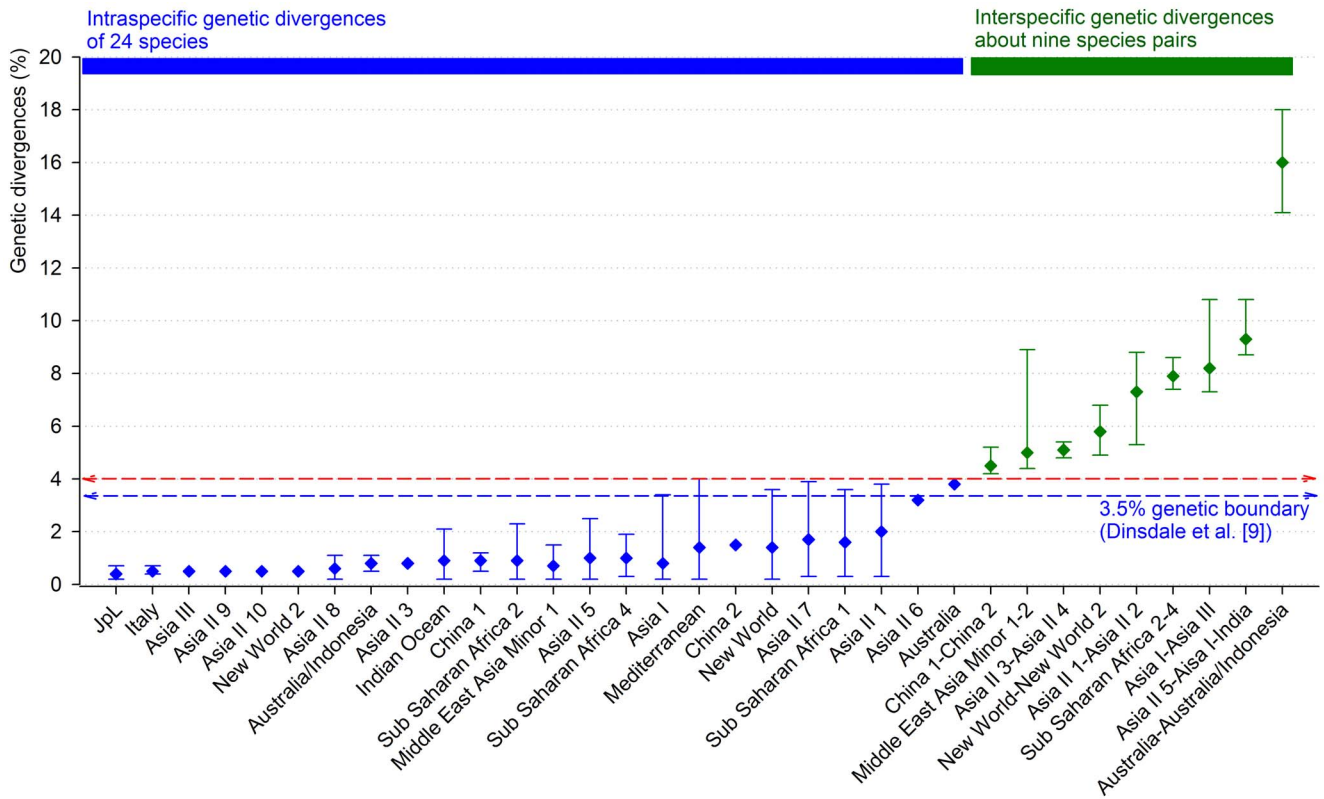


Figure 4. Comparison of intraspecific genetic divergences in 24 species and interspecific genetic divergences in nine species pairs.
doi:10.1371/journal.pone.0063817.g004

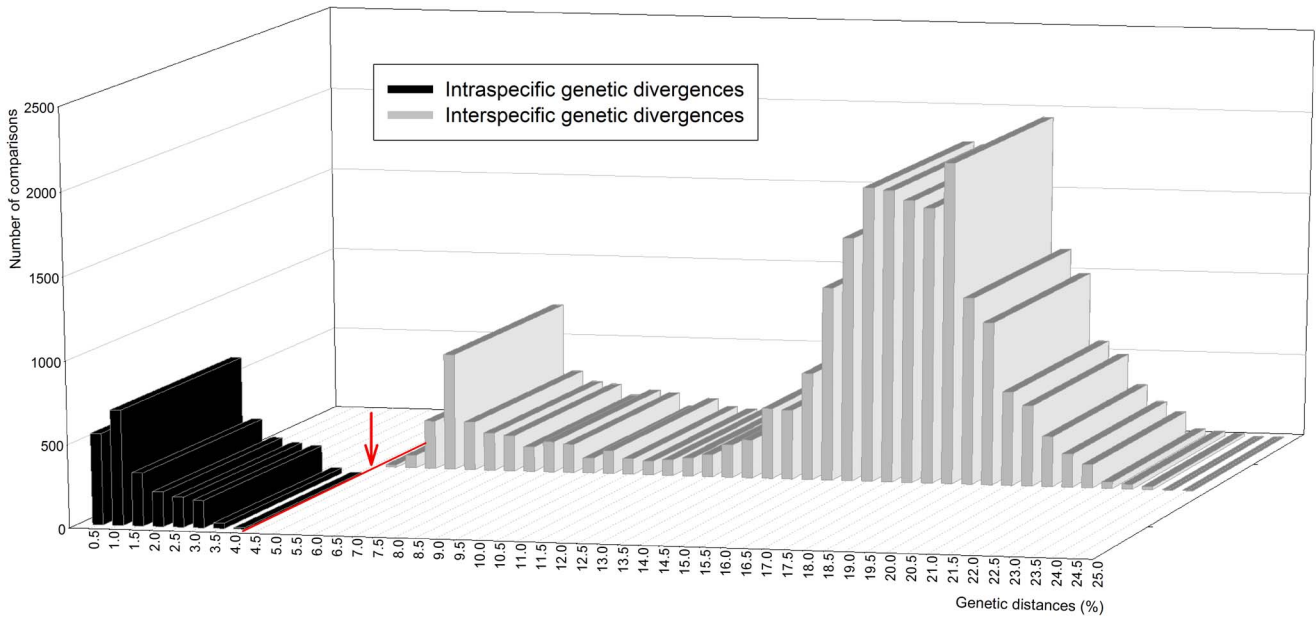


Figure 5. Distribution of genetic divergences based on the Kimura-2 parameter analysis for *COI* sequences according to two taxonomic levels.
doi:10.1371/journal.pone.0063817.g005

were conducted using both CLUSTALW 1.83 [29], with default parameters, and a BLAST search [30] in the Insect Mitochondrial Genome Database (IMGD; <http://www.imgd.or/>) [31]. A complete *COI* sequence of *B. tabaci* (NC 006279) was used as the standard sequence in the alignments.

A total of 1568 *COI* sequences belonging to 154 species were selected based on the alignment results. The 1568 *COI* sequences consisted of 1059 *COI* sequences of *B. tabaci* and 509 *COI* sequences of 153 hemipteran species belonging to 53 genera and 8 families (Table S4).

Comparing Genetic Divergence between *B. tabaci* and Other Hemipteran Species

Firstly, three datasets were constructed from the 1568 *COI* sequences of *B. tabaci* and the 153 hemipteran species: i) *B. tabaci* dataset consisting of 1059 *COI* sequences of *B. tabaci*, ii) hemipteran species dataset consisting of 509 *COI* sequences of the 153 hemipteran species, and iii) combined dataset consisting of

Table 3. Wilcoxon signed rank tests of interspecific difference among loci.

Locus pairs		Relative Ranks, <i>n</i> , <i>P</i> -value	Result
W+	W-		
<i>COI</i>	<i>ATP6</i>	W+ = 0, W- = 96, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>ATP6</i>
<i>COI</i>	<i>COII</i>	W+ = 1, W- = 96, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>COII</i>
<i>COI</i>	<i>CytB</i>	W+ = 10, W- = 99, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>CytB</i>
<i>COI</i>	<i>ND1</i>	W+ = 11, W- = 99, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>ND1</i>
<i>COI</i>	<i>ND2</i>	W+ = 4, W- = 97, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>ND2</i>
<i>COI</i>	<i>ND3</i>	W+ = 0, W- = 96, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>ND3</i>
<i>COI</i>	<i>ND4</i>	W+ = 0, W- = 96, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>ND4</i>
<i>COI</i>	<i>ND5</i>	W+ = 0, W- = 96, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> < <i>ND5</i>
<i>COI</i>	<i>IrRNA</i>	W+ = 98, W- = 16, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> > <i>IrRNA</i>
<i>COI</i>	<i>srRNA</i>	W+ = 102, W- = 24, n = 190, <i>P</i> ≤ 0.000	<i>COI</i> > <i>srRNA</i>

n is the number of comparison pairs. *P*-value is one sided probability of divergence rates being equal. *P*-values less than 0.05 were considered significant and interpreted to reflect significant differences in observed rates of divergences.

doi:10.1371/journal.pone.0063817.t003

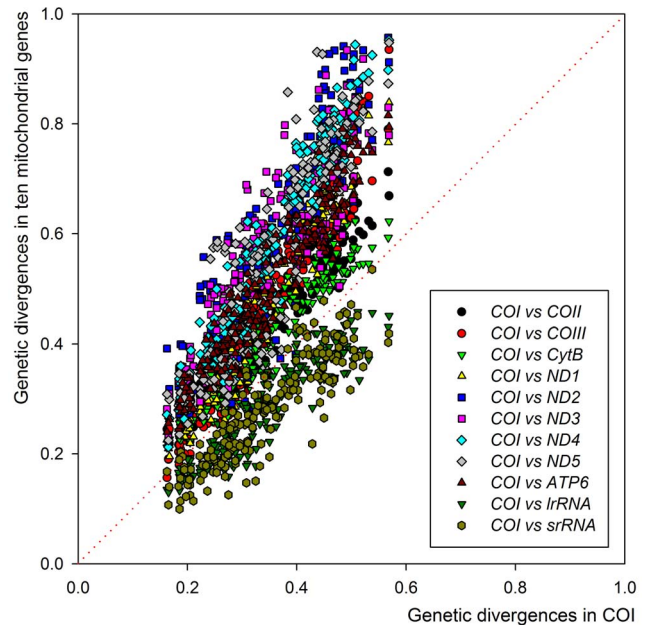


Figure 6. Relationships between the genetic distance in *COI* and that in another mitochondrial gene (*ATP6*, *COII*, *CytB*, *ND1*, *ND2*, *ND3*, *ND4*, *ND5*, *IrRNA*, and *srRNA*) for the same pair of hemipteran species.
doi:10.1371/journal.pone.0063817.g006

1568 *COI* sequences of the 153 hemipteran species+*B. tabaci*. The alignments of *COI* sequences of the three datasets were conducted by using both CLUSTALW 1.83 [29], with default parameters, and a BLAST search [30] in the IMGD [31].

Genetic divergence values for the three datasets were calculated using the distance matrix analysis function implemented in the IMGD. Genetic divergence of the *B. tabaci* dataset was calculated at one taxonomic level, among conspecific individuals, whereas those of the remaining datasets were analyzed according to three taxonomic levels, 1) among conspecific individuals, 2) among congeneric species, and 3) among confamilial species.

Defining the Number of Species of the *B. tabaci* Complex

To construct a phylogenetic tree, we first conducted neighbor-joining (NJ) analyses based on the 1059 *COI* sequences of *B. tabaci* using MEGA 5.0 [32] and excluded duplicate haplotypes in these sequences based on the NJ tree. The duplicate haplotypes were rechecked using TCS [33]. In total, 212 *COI* sequences (20% out of the 1059 *COI* sequences) were selected.

Phylogenetic trees were constructed based on the 212 *COI* sequences of *B. tabaci* and ten *COI* sequences of *Bemisia atriplex*, *Bemisia subdecipiens*, *Bemisia afer*, *Aleurocanthus camelliae*, *Aleurodicus dispersus*, *Aleurocanthus spiniferus*, *Trialeurodes lauri*, *Trialeurodes ricini*, and *Trialeurodes vaporariorum* as one outgroup. Alignments of nucleotide sequences were performed using CLUSTALX with default conditions [29]. The well-aligned blocks from the nucleotide sequences were then selected with GBlocks 0.91b, again under default conditions [34]. Phylogenetic reconstruction was done using Bayesian inference (BI) analysis. The BI trees were obtained using MrBayes 3.1.2 [35] and the Markov Chain Monte Carlo technique (MCMC). We chose the GTR+G model for the dataset based on the hierarchical likelihood ratio test that was conducted using MrModeltest v2.2 [36]. Model parameter values were treated as unknowns and were estimated for each analysis. Random starting trees were used, and analyses were run for 10000000 generations with sampling done every 100 generations for the *B. tabaci* dataset. Bayesian posterior probabilities were then calculated from the sample points after the MCMC algorithm began to converge. To ensure that our analyses are not trapped in local optima, four independent MCMC runs were performed. Topologies and posterior clade probabilities from different runs were compared for congruence.

Because of the large number of *COI* sequences of *B. tabaci*, a canonical discriminant analysis (CDA), as implemented in SPSS (IBM® SPSS Statistics, Ver 20, IBM), was used to demonstrate a graphical summary of the species-grouping results. CDA explores differences in species assemblage compositions by using non-parametric permutation tests which provide exact probability values, and the similarity or dissimilarity of the various sampling units are graphically displayed. A matrix was constructed based on 22336 pairwise genetic distances from the 212 *COI* sequences using on the Kimura 2-parameter (K2P) distance model [37] in MEGA 5.0 [32].

Comparing Intraspecific and Interspecific Genetic Divergences

The 212 *COI* sequences were categorized into several species based on the phylogenetic tree (see Results). We tentatively named a total of 31 species based on five previous studies: *Asia I*, *Asia II 1*, *Asia II 2*, *Asia II 3*, *Asia II 4*, *Asia II 5*, *Asia II 6*, *Asia II 7*, *Asia II 8*, *Australia*, *Australia/Indonesia*, *China 1*, *China 2*, *Italy*, *Mediterranean*, *Middle East Asia Minor 1*, *Middle East Asia Minor 2*, *Indian Ocean*, *New World*, *Sub Saharan Africa 1*, *Sub Saharan Africa 2*, *Sub Saharan Africa 3*, *Sub Saharan Africa 4*, and *Uganda* [9], *JpL* [15], *Asia I-India* [11], *Asia*

III, *China 3*, *Asia II 9*, *Asia II 10* [13], and *New World 2* [12]. Of these species, seven species, *Asia I-India*, *Asia II 2*, *Asia II 4*, *China 3*, *Middle East Asia Minor 2*, *Sub Saharan Africa 3*, and *Uganda*, were excluded in this analysis because each group included only a single haplotype. Intraspecific genetic divergences were analyzed based on the remaining 24 species.

Interspecific genetic divergence was calculated between pairs of closely related species. We selected nine pairs of closely related species consisting of ten species based on the phylogenetic tree (see Results): *Asia II 1-Asia II 2*, *Asia I-Asia III*, *Asia II 3-Asia II 4*, *Asia II 5-Asia I-India*, *Australia/Indonesia-Australia*, *China 1-China 2*, *Middle East Asia Minor 1-Middle East Asia Minor 2*, *New World-New World 2*, and *Sub Saharan Africa 2-Sub Saharan Africa 4*.

All intraspecific and interspecific genetic divergence estimates were calculated using the *p*-distance model and the K2P distance model [37] of MEGA 5.0 [32], respectively.

Analyzing Genetic Divergences of Aphid Species

A total of 47 *COI* sequences of 34 aphid species were downloaded from the Genbank (Table S5). To analyze sequences of the same length and position as in the *B. tabaci* dataset, the 47 *COI* sequences were aligned with *B. tabaci* *COI* sequences by using CLUSTALW 1.83 [29]. Genetic divergence of the aphid species was calculated according to three taxonomic levels, 1) among conspecific individuals, 2) among congeneric species, and 3) among species of different genera in the same family, by using the K2P distance model [37] of MEGA 5.0 [32].

Comparing Genetic Divergences of Eleven Mitochondrial Genes

To compare the pattern of genetic divergences among genes, we chose eleven mitochondrial genes, ATP synthase F0 subunits 6 (ATP6), COI, cytochrome oxidase subunits II (COII), cytochrome oxidase subunits III (COIII), NADH dehydrogenase subunits 1–5 (ND1, ND2, ND3, ND4, and ND5), large subunit ribosomal RNA (rRNA), and small subunit ribosomal RNA (srRNA), which have frequently been used in insect molecular studies. A total of 20 complete mitochondrial genomes belonging to the order Hemiptera were downloaded from the Genbank (Table S6). The genetic divergences of the eleven genes were calculated according to two taxonomic levels (between genera within the same suborder and between suborder within the same order) by using the K2P distance model [37] of MEGA 5.0 [32]. Wilcoxon signed rank tests were performed to compare interspecific variability between COI and other ten genes following Kress & Erickson [38].

Supporting Information

Table S1 Intraspecific generic divergences about 31 species of the *Bemisia tabaci* complex.

(DOC)

Table S2 Intraspecific generic divergences between nine species pairs.

(DOC)

Table S3 Genetic divergences of 11 mitochondrial genes in 20 hemipteran complete mitochondrial genomes.

(DOC)

Table S4 The list of 1059 individuals of *Bemisia tabaci* and 509 individuals of 153 hemipteran species, 53 genera, and 8 subfamilies.

(DOC)

Table S5 The list of 47 individuals of 34 aphid species.

(DOC)

Table S6 The list of 20 hemiptera mitochondrial genomes.

(DOC)

References

- De Barro PJ, Driver F, Trueman JW, Curran J (2000) Phylogenetic relationships of world populations of *Bemisia tabaci* (Gennadius) using ribosomal ITS1. *Mol Phylogenet Evol* 16: 29–36.
- Brown J, Frohlich D, Rosell R (1995) The sweetpotato or silverleaf whiteflies: biotypes of *Bemisia tabaci* or a species complex? *Annu Rev Entomol* 40: 511–534.
- De Barro PJ, Liu SS, Boykin LM, Dinsdale AB (2011) *Bemisia tabaci*: a statement of species status. *Annu Rev Entomol* 56: 1–19.
- Byrne DN, Bellows TS (1991) Whitefly biology. *Ann Rev Entomol* 36: 431–457.
- Brown JK (1990) An update on the whitefly-transmitted geminiviruses in the Americas and the Caribbean Basin. *FAO Plant Prot Bull* 39: 5–23.
- Brown JK (2000) Molecular markers for the identification and global tracking of whitefly vector-Begomovirus complexes. *Virus Res* 71: 233–260.
- Costa HS, Brown JK (1991) Variation in biological characteristics and in esterase patterns among populations of *Bemisia tabaci* (Genn) and the association of one population with silverleaf symptom development. *Entomol Exp Appl* 61: 211–219.
- Bedford ID, Briddon RW, Brown JK, Rosell RC, Markham PG (1994) Geminivirus-transmission and biological characterisation of *Bemisia tabaci* (Gennadius) biotypes from different geographic regions. *Ann Appl Biol* 125: 311–325.
- Dinsdale A, Cook L, Riginos C, Buckley YM, De Barro P (2010) Refined Global Analysis of *Bemisia tabaci* (Hemiptera: Sternorrhyncha: Aleyrodidae: Aleyrodidae) Mitochondrial Cytochrome Oxidase 1 to Identify Species Level Genetic Boundaries. *Ann Entomol Soc Am* 103: 196–208.
- Campbell BC, Duffus JE, Baumann P (1993) Determining whitefly species. *Science* 261: 1333; author reply 1334–1335.
- Chowda-Reddy RV, Kirankumar M, Seal SE, Muniyappa V, Valand GB, et al. (2012) *Bemisia tabaci* Phylogenetic Groups in India and the Relative Transmission Efficacy of Tomato leaf curl Bangalore virus by an Indigenous and an Exotic Population. *J Integr Agric* 11: 235–248.
- Alemandri V, De Barro P, Bejerman N, Arguello Caro EB, Dumon AD, et al. (2012) Species within the *Bemisia tabaci* (Hemiptera: Aleyrodidae) complex in soybean and bean crops in Argentina. *J Econ Entomol* 105: 48–53.
- Hu J, De Barro P, Zhao H, Wang J, Nardi F, et al. (2011) An Extensive Field Survey Combined with a Phylogenetic Analysis Reveals Rapid and Widespread Invasion of Two Alien Whiteflies in China. *PLoS ONE* 6(1): e16061. doi:10.1371/journal.pone.0016061.
- Lee W, Kim H, Lim J, Choi H-r, Kim Y, et al. (2011) Barcoding aphids (Hemiptera: Aphididae) of the Korean Peninsula: updating the global data set. *Mol Ecol Resour* 11: 32–37.
- Ueda S, Kitamura T, Kijima K, Honda K-I, Kanmiya K (2008) Distribution and molecular characterization of distinct Asian populations of *Bemisia tabaci* (Hemiptera: Aleyrodidae) in Japan. *J Appl Entomol* 133: 355–366.
- Howard DJ, Berlocher SH (1998) *Endless Forms: Species and Speciation*. Oxford University Press, New York.
- Sites JW, Marshall JC (2003) Delimiting species: a Renaissance issue in systematic biology. *Trends Ecol Evol* 18: 462–470.
- Mayr E (1940) Speciation phenomena in birds. *Am Nat* 74: 249–278.
- De Barro P, Ahmed MZ (2011) Genetic Networking of the *Bemisia tabaci* Cryptic Species Complex Reveals Pattern of Biological Invasions. *PLoS ONE* 6(10): e25579. doi:10.1371/journal.pone.0025579.
- Hebert PD, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proc Biol Sci* 270: 313–321.
- Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PD (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proc Natl Acad Sci U S A* 103: 968–971.
- Smith MA, Woodley NE, Janzen DH, Hallwachs W, Hebert PD (2006) DNA barcodes reveal cryptic host-specificity within the presumed polyphagous members of a genus of parasitoid flies (Diptera: Tachinidae). *Proc Natl Acad Sci U S A* 103: 3657–3662.
- Smith MA, Rodriguez JJ, Whitfield JB, Deans AR, Janzen DH, et al. (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc Natl Acad Sci U S A* 105: 12359–12364.
- Brower A (1994) Phylogeny of Heliconius butterflies inferred from mitochondrial DNA sequences (Lepidoptera: Nymphalidae). *Mol Phylogenet Evol* 3: 159–174.
- Lee W, Koh S-H, Choi WI, Jung CS, Kim I-K, et al. (2012) Barcoding forest insect pests in South Korea: Constructing a basic endemic species dataset. *J Asia Pac Entomol* 15: 363–368.
- Footitt RG, Maw HEL, Dohlen CDV, Herbert PDN (2008) Species identification of aphids (Insecta: Hemiptera: Aphididae) through DNA barcodes. *Mol Ecol Resour* 8: 1189–1201.
- Calvert LA, Cuervo M, Arroyave JA, Constantino LM, Bellotti A, et al. (2001) Morphological and Mitochondrial DNA Marker Analyses of Whiteflies (Homoptera: Aleyrodidae) Colonizing Cassava and Beans in Colombia. *Ann Entomol Soc Am* 94: 512–519.
- Dubey AK, Ko C-C, David BV (2009) The Genus *Lipaleyrodes* Takahashi, a Junior Synonym of *Bemisia* Quaintance and Baker (Hemiptera: Aleyrodidae): A Revision Based on Morphology. *Zool Stud* 48: 539–557.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25: 4876–4882.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
- Lee W, Park J, Choi J, Jung K, Park B, et al. (2009) IMGd: an integrated platform supporting comparative genomics and phylogenetics of insect mitochondrial genomes. *BMC Genomics* 10: 148.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731–2739.
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9: 1657–1659.
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17: 540–552.
- Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- Nylander JAA (2004) MrModeltest v2.2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16: 111–120.
- Kress WJ, Erickson DL (2007) A Two-Locus Global DNA Barcode for Land Plants: The Coding rbcL Gene Complements the Non-Coding trnH-psbA Spacer Region. *PLoS ONE* 2(6): e508. doi:10.1371/journal.pone.0000508.

Author Contributions

Conceived and designed the experiments: WL JP GSL SA. Performed the experiments: WL. Analyzed the data: WL JP. Contributed reagents/materials/analysis tools: WL JP GSL SL SA. Wrote the paper: WL JP GSL SL SA.